AN EMBEDDED REED-SOLOMON CODING AND QAM SCHEME FOR SUB-BAND CODED SPEECH TRANSMISSION VIA RAYLEIGH-FADING CHANNELS

L. Hanzo, R. Steele, P.M.Fortune; University of Southampton SO9 5NH

1. Introduction

With the incentive of accommodating the highest possible number of speech transmission channels in the bandwidth available for the digital cellular mobile radio network of the near future we embark on the investigation of a spectrally efficient, jointly optimised sub-band codec (SBC), Reed-Solomon (RS) codec, 16-level quadrature amplitude modulator (16-QAM) scheme.

2. Bit-Error Rate Performance of 16-QAM Channels

2.1. Signal constellation

In 16-QAM systems one state of the 16-point signal constellation depicted in Fig.1 is represented by four Gray-coded input bits. Hence whenever a phasor is corrupted over the channel into one of the adjacent constellation points, only one bit out of the four is erroneous. If we denote the inphase bits by i_1, i_2 and the quadrature bits by q_1, q_2, one 16-QAM phasor is uniquely described by the four-bit word i_1, i_2, q_1, q_2. Notice that in case of i_1, q_1 half of the bits is at a distance of 3d, and half of them is at distance d in Fig.1 from the corresponding decision boundaries, while i_2, q_2 are always at distance d from the decision boundaries. Whence the bits i_1, q_1 constitute a higher integrity class one (C1) subchannel as compared to the lower integrity class two (C2) subchannel.

2.2. Performance over Gaussian Channels

According to the different 'protection distances' in Fig.1 the C1 and C2 subchannel error probabilities P_{1G}, P_{2G} are given by:

$$P_{1G} = \frac{1}{2}\left(Q\left\{ \sqrt{\frac{E}{5N}} \right\} + Q\left\{ 3\sqrt{\frac{E}{5N}} \right\} \right), \quad (1)$$

$$P_{2G} = Q\left\{ \sqrt{\frac{E}{5N}} \right\}, \quad (2)$$

where $E = 10\ d^2$ is the average energy of a phasor in Fig.1 and N is the one-sided spectral density of the additive white Gaussian noise (AWGN). The overall 16-QAM performance is given by:

$$P_G = P_{1G} + P_{2G}/2. \quad (3)$$

Simulation results depicted in Fig.2 give identical curves for P_{1G}, P_{2G} and P_G. Notice that there is a consistent but not dramatic advantage in using the C1 subchannel as compared to the C2 subchannel.

2.3. Performance over Rayleigh-Fading Channels

Closed form formulae can be derived for the performance over Rayleigh-fading channels as well under the assumption of slowly fading channels. Then the equivalent low pass received signal in one symbol interval is written as [1]:

$$r(t) = \alpha \cdot e^{-j\phi(t)} u(t) + n(t), \quad (4)$$

where α is the Rayleigh-fading envelope's attenuation, $\phi(t)$ is the received signal's phase, u(t) is the transmitted QAM signal and n(t) is the AWGN. Because of the assumption of slow fading α and $\phi(t)$ are constant for one symbol interval and the instantaneous SNR γ as well as the average SNR Γ are given by:

$$\gamma = \alpha^2 \cdot \frac{E}{N}; \tag{5}$$

$$\Gamma = \bar{\alpha}^2 \cdot \frac{E}{N}. \tag{6}$$

Then the bit error probability $P_{1R}(\Gamma)$ of the C1 subchannel at the average SNR Γ over Rayleigh-fading channels is yielded as:

$$P_{1R}(\Gamma) = \frac{1}{2\Gamma}\int_0^{\infty}\left(Q\left\{\sqrt{\frac{\gamma}{5}}\right\} + Q\left\{3\sqrt{\frac{\gamma}{5}}\right\}\right)e^{-\frac{\gamma}{\Gamma}}d\gamma. \tag{7}$$

The bit error probability $P_{2R,0}(\Gamma)$ over Rayleigh-fading channels, when incorrectly decoding a C2 logical zero into one is as follows:

$$P_{2R,0}(\Gamma) = \frac{1}{\Gamma}\int_0^{4\Gamma} Q\left\{\left(2\frac{\bar{\alpha}}{\alpha} - 1\right)\sqrt{\frac{\gamma}{5}}\right\}e^{-\frac{\gamma}{\Gamma}}d\gamma + \frac{1}{\Gamma}\int_{4\Gamma}^{\infty}\left(1 - Q\left\{\left(1 - 2\frac{\bar{\alpha}}{\alpha}\right)\sqrt{\frac{\gamma}{5}}\right\}\right)e^{-\frac{\gamma}{\Gamma}}d\gamma, \tag{8}$$

while the probability of incorrectly decoding a C2 logical one into zero over Rayleigh-fading channels is as below:

$$P_{2R,1}(\Gamma) = \frac{1}{\Gamma}\int_0^{4\Gamma/9}\left(1 - Q\left\{\left(2 - 3\frac{\bar{\alpha}}{\alpha}\right)\sqrt{\frac{\gamma}{5}}\right\}\right)e^{-\frac{\gamma}{\Gamma}}d\gamma + \frac{1}{\Gamma}\int_{4\Gamma/9}^{\infty} Q\left\{\left(3\frac{\bar{\alpha}}{\alpha} - 2\right)\sqrt{\frac{\gamma}{5}}\right\}e^{-\frac{\gamma}{\Gamma}}d\gamma. \tag{9}$$

Hence, in case of a random sequence, when zero and one are equiprobable, the C2 error probability $P_{2R}(\Gamma)$ over Rayleigh-fading channels is given by:

$$P_{2R}(\Gamma) = \frac{1}{2}\left\{P_{2R,0}(\Gamma) + P_{2R,1}(\Gamma)\right\}. \tag{10}$$

The curves given by Eq.(7) and Eq.(10) are also depicted in Fig.2 and are verified by simulation results, obtained using a Rayleigh-fading envelope sampled at 16 ksample/s, where the propagation frequency is 900 MHz. In contrast to the AWGN channel, the advantage in using the C1 subchannel is dramatic, resulting into approximately 1 % BER at SNR = 20 dB, which by using RS coding allows SBC speech communication, while the C2 subchannel is unusable.

3. 16-QAM Performance Enhancement

3.1. Automatic Gain Control Methods

The 16-QAM system performance is radically improved over Rayleigh-fading channels if some degree of adaptive fade-tracking or attenuation compensation is deployed. One method implemented is a kind of channel sounding, in which we regularly transmit a sounding symbol of maximum amplitude, e.g. phasor P_4 in Fig.1. The received phasor P_4 is then used to derive a gain factor for the consecutive k number of phasors to compensate for the fading attenuation. On the price of a slight transmission overhead this reduces the C2 BER.

However, an equally effective automatic gain control (AGC) method can be derived without transmission overhead by comparing the average energy R of the previous k received phasors to the theoretical expected average energy $E=10\ d^2$, which yields a scaling factor G for the k received phasors: G=E/R. We refer to this method as average locking AGC system.

The 16-QAM BER performance of C1 and C2 subchannels was experimentally investigated for various channel SNRs and vehicular speeds over a range of k values. The C1 subchannel's performance is not significantly altered by the AGC, but that of the C2 subchannel is dramatically improved, as evidenced by Fig.3. In the simulations for each Rayleigh-fading envelope sample one QAM sample is transmitted, ie the transmission rate is 16 kBd. Hence the 120 mph curve in Fig.3 corresponds to 60 mph at 8 kBd and to 30 mph at 4 kBd. Since we are interested in systems with transmission rates between 4 kBd and 8 kBd, the curves corresponding to higher vehicle speeds represent more important cases for our investigations. Observe that nearly all curves, ie conditions have an optimum k value, and the AGC fade-tracking capability drops with increasing k. Hence for further investigations k=4 is selected as a good compromise for all conditions.

3.2. 16-QAM Constellation Distortion

A plausible way of equalising the BER performance of C1 and C2 subchannels is to distort the original signal constellation by moving the C1 constellation phasors nearer to their respective decision boundaries and ensuring higher protection distance for the C2 subchannel phasors. We define the constellation distortion factor CD as:

$$CD = 20\,lg\,\frac{d}{d-d_1}[dB], \tag{11}$$

where d_1 represents the protection distance reduction, caused by shifting the C1 constellation points symmetrically towards both of their decision boundaries. The subchannel BERs at SNR = 20 dB are depicted in Fig.4 as a function of CD, where we observe that around CD=3.5 dB the two subchannels have identical BERs, but their average is also slightly increased to 2.5 %. This way we arrive at a system with two identical subchannels, which is still unsuited for SBC speech communication, but the resulted BER=2.5 % is now low enough to be combated by RS coding.

3.3. Diversity Techniques

Our simulation results with d-th order diversity confirm the effectiveness of the method in combating the fading of the channel. Each of the d received signal is generated by modifying the transmitted phasor by the Rayleigh-fading envelope, using a different, random start point. Also a separate AWGN source was used for each incoming signal. Each of the d received signals is assigned a separate average locking AGC system, and at each sampling instant the channel exhibiting the highest average signal level over the interval k=4 is selected for 16-QAM demodulation.

The 16-QAM C1 and C2 subchannel logarithmic BER is displayed in Fig.5 for the SNR values of 20 dB, 22 dB and 25 dB, respectively, as a function of the diversity order. While the C2 subchannel BER performance is hardly effected by deploying diversity, the C1 subchannel performance is cut by two orders of magnitudes, if the diversity order is d=4. Even in case of the more practical second order diversity the C1 subchannel BER is decreased below 0.1 % at

SNR = 20 dB. Hence, the C1 subchannel is well suited for SBC speech transmission, while the BER=2 % of the C2 subchannel can be tackled by RS coding to allow speech communication.

4. Proposed SBC/RS/16-QAM Systems

Based on our previous studies [2] and present deductions, a number of combined SBC/RS/16-QAM systems with various speech qualities and transmission rates can now be contrived. As an objective quality measure for system comparison the SBC segmental SNR (SEGSNR) is used, and for our SBC implementation a SEGSNR=11 dB is associated with near-toll quality subjective assessment.

Since the SBC codec has a transmission rate of 16 kbit/s, if no RS coding is deployed, the resultant transmission speed via 16-QAM is 16 kbit/s/4 = 4 kBd. However, this system is not capable of delivering near-toll quality speech, as evidenced by the SEGSNR curve in Fig.6. This system is merely used as a bench-marker for our candidate schemes.

Another system can be devised by equalising the BERs of both subchannels by the help of constellation distortion with CD=3.5 dB, but because of the inherently high BER at least a half-rate RS code has to be utilised. This renders the system to be not competitive because of its high 8 kBd transmission rate, hence this system concept is dropped as well.

In the following two systems we decided to actually exploit the BER difference between the subchannels. In the first system we deploy RS coding with different coding rates for the two subchannels to equalise the BER differences after RS decoding. In the second system we use the subchannels as two independent systems, and arrange for the perceptually and objectively important SBC speech bits to be transmitted over the better subchannel.

4.1. The Sub-band Codec

The sub-band codec [3] produces eight sub-bands each of 500 Hz bandwidth by means of a quadrature mirror filter (QMF) bank. Each sub-band signal is encoded using Jayant's one-word memory quantiser whose number of bits is either 0, 2, 3 or 4, depending on which signal classification is used. The classification, given in the first column in Table 1, is determined by the distribution of spectral energy in the speech signal. Each input speech sample results in 15 bits being produced from the sub-band quantisers and yields a rate of 15 kb/s as each sub-band is sampled at 1 kHz. The bit allocation is reviewed every 6 ms, and the classification adopted is represented by a 2-bit word that is repeated twice for channel protection. Thus a SBC frame consists of (6x15) + 6 = 96 bits with a duration of 6 ms.

4.2. The 6.7 kBaud System

The philosophy behind this system is to add channel coding to both subchannels in such a way that the BER of each is not only low enough to carry toll-quality SBC speech, but also nearly matched. As a consequence no special mapping strategies are required. In the absence of RS coding, but in the presence of average locking AGC, the BER for the C2 subchannel is approximately three times that encountered in the C1 subchannel. Therefore we select the RS code for the C2 subchannel to have three times the error correcting capability of the RS code used for the C1 subchannel. Because the total protected data rate passing via each of the subchannels must be equal, the redundancy of the code used with the C1 subchannel must be matched to that used in the C2 subchannel. Furthermore, it is convenient to arrange that a single RS coded frame be composed of an integer number of SBC frames. The channel codes selected to meet these criteria were RS(120,96) for the C1 subchannel and RS(120,48) for the C2, both having 8-bit symbols. Thus bits in successive SBC frames were sorted into two contiguous streams by placing every third bit in stream A and the others in stream B. For every 12 frames, ie 1152 bits, 768 bits in stream A were RS(120, 96) coded to yield the C1 subchannel data. The C2 subchannel data was composed of 384 bits from stream B that were RS(120, 48) coded.

The segmental SNR of the speech signal as a function of E/N is shown in Fig.6 for a vehicular speed of 30 mph. Near toll-quality speech can be obtained when the channel SNR is in excess of approximately 23 dB.

4.3. The 5.3 kBaud System

Even with the application of AGC the BER of the C2 subchannel remains unacceptably high. By contrast the C1 QAM data subchannel has a BER that is sufficiently low for transmission of the least significant bits (LSB) in the SBC data frame. By RS coding the bits passing through the C2 subchannel the BER can be decreased below that of the unprotected C1 subchannel. However, the effect of RS coding also results in an increase in the bit rate, and as the C1 and C2 subchannels must each have two bits in each Gray coded constellation point, it follows that the information throughput on the C2 subchannel is reduced. It is essential to RS code the classifier bits, and it is desirable that the most significant bit (MSB) of each coded sub-band signal in the SBC frame be RS coded. However, there are too many sub-band MSBs to be RS coded. For example, in the presence of either an intermediate or unvoiced SBC frame, 66 bits in each 96-bit frame would need to be error protected. The weakest RS code with sufficient error correction capability to improve the BER of the C2 subchannel to an acceptable level is the RS(120,60) code operating in 'Galois Field' GF(256). This code fixes the data capacity of the C2 subchannel at 1/2 of that of the C1 subchannel allowing 32 bits of each 96-bit frame to be transmitted via the higher performance route. The resultant data throughput is 21.3 kbit/s for 16 kbit/s speech. This is equivalent to a symbol rate of 5.3 kBaud using the 16-level QAM system.

As erroneous reception of the classifier information bits can cause severe degradation of the recovered speech they are always transmitted via the RS coded C2 subchannel. The MSBs of the code words from up to four sub-bands are also mapped onto C2 locations in order to optimise performance. These sub-bands are chosen to be those where the main proportion of the speech energy is likely to be distributed. The mappings are displayed in Table 1. For the voiced speech classification we see that the solid vertical lines separate the sub-bands which increase in frequency from left to right as 0 - 0.5, 0.5 - 1.0, 1.1 - 1.5, 1.5 - 2.0 and 2.0 - 2.5 kHz. Also see Table 1. The number of columns within each sub-band represents the number of bits used to quantise each sub-band signal. Each row has 15 bits representing the SBC output for each input speech sample. The six rows constitute the SBC output for the frame, although the first bits 1 to 6 in the frame (representing the classification) are not shown. The numbers in Table 1 indicate the bit position in the frame, where bit 96 is the last in the frame. We see that for voiced speech the LSBs for the sub-band 0.5 to 1.0 kHz are 11, 26, 41, 56, 71, 86, while the MSBs are 14, 29, 44, 59, 74 and 89. The bits assembled for RS coding are those encircled.

In the simulations two speech sentences were SBC encoded. For every frame the appropriate sub-table (a), (b), (c) or (d) in Table

1. was used and the encircled bits RS(120, 60) coded. The C1 and the RS coded C2 sub-channels were transmitted via 16-level QAM over a Rayleigh-fading channel. Upon demodulation, RS decoding, and SBC decoding, the speech signal was recovered and the segmental SNR computed. The variation of segmental SNR as a function of channel SNR (E/N) is displayed in Fig.6 for a vehicular speed of 30 mph. For near toll-quality speech the channel SNR should exceed 30 dB. Although we have not displayed in Fig.6, our experiments demonstrate that less than 1.5 dB separates the channel SNRs when the vehicular speed was doubled and quadrupled.

The performance of the 5.3 kBd system without diversity is slightly more modest, than that of the 6.7 kBd scheme. However, by deploying second order diversity the BER of the C1 subchannel is dramatically improved, as we have seen in Sec.3. Since 64 bits out of each 96 bits frame are transmitted via this high quality subchannel, a highly extended SNR operation range is achieved, as demonstrated by the best SEGSNR performance in Fig.6.

5. Conclusion

A 5.3 kBd combined SBC/RS/16-QAM system with second order diversity and average locking fade-tracking capability yields near-toll quality speech for channel SNRs above 20-22 dB within a bandwidth of 10 kHz. Further subjective quality improvement is anticipated, when the system is optimised for the regular pulse excitation codec with long term predictor (RPE-LTP), while enhanced bandwidth efficiency can be achieved by deploying a code excited linear predictive (CELP) speech codec.

References

[1] J.G. Proakis: Digital communications, McGraw-Hill, 1983

[2] L. Hanzo, K.H.J. Wong, and R. Steele, "Jointly optimised sub-band coding and channel coding techniques for mobile communications", accepted by IEEE Trans. on Communications.

[3] R.B. Hanes: A 16 kbit/s speech codec for four channel 64 kbit/s transmission, British Telecom. Techn. Journ., Vol. 3., No.1 Jan 1985

5.3k baud system

The bits for RS coding in the SBC frame

CLASSIFICATION	SUB-BANDS IN kHz				
	0.0 - 0.5	0.5 - 1.0	1.0 - 1.5	1.5 - 2.0	2.0 - 2.5
VOICED (a)	7 8 9 ⑩	11 12 13 ⑭	15 16 ⑰	18 ⑲	20 ㉑
	22 23 24 ㉕	26 27 28 ㉙	30 31 ㉜	33 ㉞	35 36
	37 38 39 ㊵	41 42 43 ㊹	45 46 ㊼	48 ㊾	50 51
	52 53 54 (55)	56 57 58 (59)	60 61 (62)	63 (64)	65 (66)
	67 68 69 (70)	71 72 73 (74)	75 76 (77)	78 (79)	80 81
	82 83 84 (85)	86 87 88 (89)	90 91 (92)	93 (94)	95 96
	ONLY CIRCLED NUMBERS LISTED IN SUB-TABLES (b) - (d)				
INTERMEDIATE (b)	10 13 15 17 19 25 28 30 32 40 43 45 47 55 58 60 62 64 70 73 75 77 85 88 90 92				
UNVOICED (c)	8 11 14 17 19 26 29 32 38 41 44 47 49 53 56 59 62 64 71 74 77 83 86 89 92 94				
VOICE BAND DATA (d)	10 13 14 18 21 25 29 33 36 40 44 48 51 55 58 59 63 66 70 74 78 81 85 89 93 96				

Table 1. SBC bit-allocation and embedded RS coding scheme

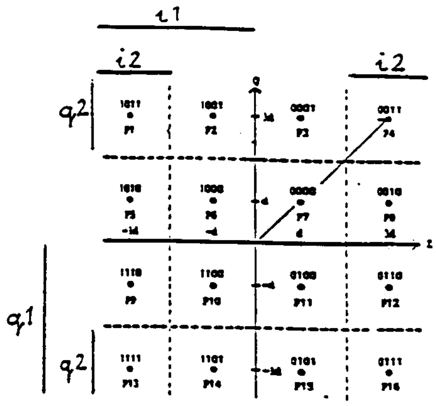

Fig.1. 16-level QAM constellation.

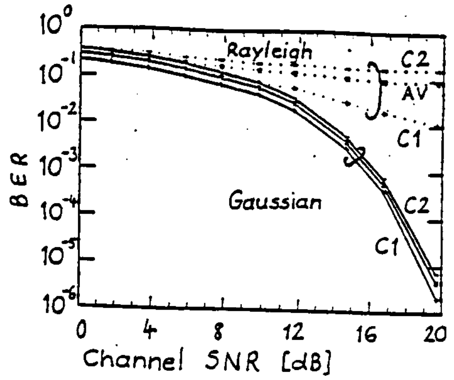

Fig.2. C1 and C2 subchannel BER and average BER versus channel SNR over AWGN and Rayleigh-fading channels

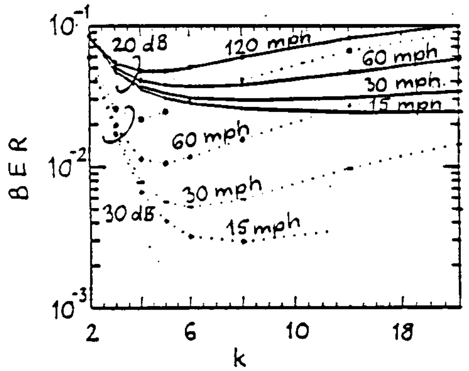

Fig.3. BER versus k for C2 subchannel with average locking over Rayleigh-fading channel

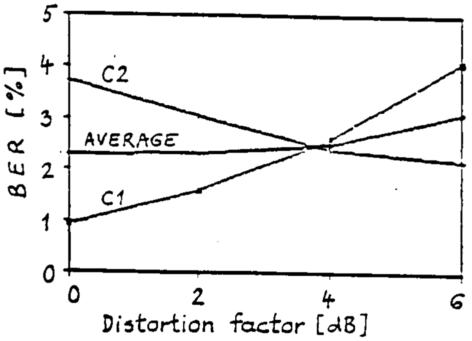

Fig.4. BER versus constellation distortion factor [dB] of C1 and C2 subchannel and that of their average

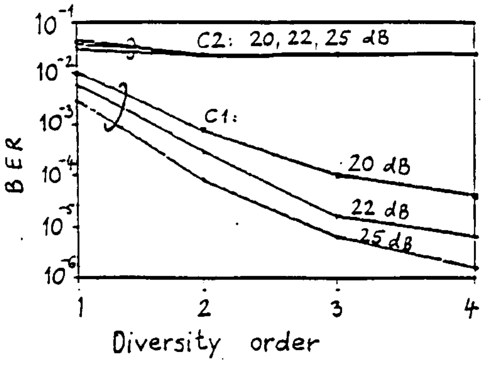

Fig.5. BER versus diversity order for C1 and C2 subchannels at SNR = 20,22,25 [dB]

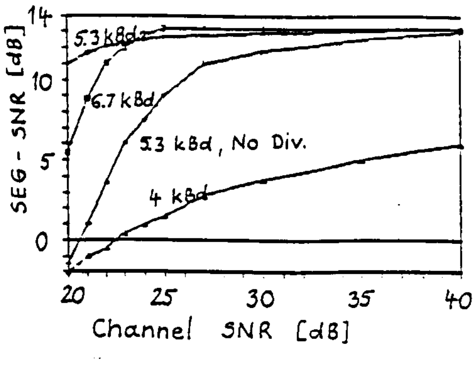

Fig.6. SEGSNR of various SBC/RS/QAM systems versus channel SNR over Rayleigh-fading channel